

S. Sakel/ARIS

RE-RUN



1600

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/883,152

DATE: 04/08/2003

TIME: 14:47:35

Input Set : A:\Pto.amc

Output Set: N:\CRF4\04082003\I883152.raw

```

4 <110> APPLICANT: Kennedy, Giulia
5      Kang, Sanmao
6      Reinhard, Christoph
7      Jefferson, Anne Bennett
9 <120> TITLE OF INVENTION: POLYNUCLEOTIDES RELATED TO COLON CANCER
12 <130> FILE REFERENCE: 2300-1663
C--> 14 <140> CURRENT APPLICATION NUMBER: US/09/883,152
C--> 15 <141> CURRENT FILING DATE: 2001-06-15
17 <150> PRIOR APPLICATION NUMBER: 60/211,835
18 <151> PRIOR FILING DATE: 2000-06-15
20 <160> NUMBER OF SEQ ID NOS: 127
22 <170> SOFTWARE: FastSEQ for Windows Version 4.0
24 <210> SEQ ID NO: 1
25 <211> LENGTH: 564
26 <212> TYPE: DNA
27 <213> ORGANISM: H. sapiens
29 <220> FEATURE:
30 <221> NAME/KEY: CDS
31 <222> LOCATION: (21)...(396)
33 <400> SEQUENCE: 1
34 ggcgagattt gtgcggcgac atg aaa ctg ctt acc cac aat ctg ctg agc tcg      53
35           Met Lys Leu Leu Thr His Asn Leu Leu Ser Ser
36           1           5           10
38 cat gtg cgg ggg gtg ggg tcc cgt ggc ttc ccc ctg cgc ctc cag gcc      101
39 His Val Arg Gly Val Gly Ser Arg Gly Phe Pro Leu Arg Leu Gln Ala
40           15           20           25
42 acc gag gtc cgt atc tgc cct gtg gaa ttc aac ccc aac ttc gtg gcg      149
43 Thr Glu Val Arg Ile Cys Pro Val Glu Phe Asn Pro Asn Phe Val Ala
44           30           35           40
46 cgt atg ata cct aaa gtg gag tgg tcg gcg ttc ctg gag gcg gcc gat      197
47 Arg Met Ile Pro Lys Val Glu Trp Ser Ala Phe Leu Glu Ala Ala Asp
48           45           50           55
50 aac ttg cgt ctg atc cag gtg ccg aaa ggg ccg gtt gag gga tat gag      245
51 Asn Leu Arg Leu Ile Gln Val Pro Lys Gly Pro Val Glu Gly Tyr Glu
52 60           65           70           75
54 gag aat gag gag ttt ctg agg acc atg cac cac ctg ctg ctg gag gtg      293
55 Glu Asn Glu Glu Phe Leu Arg Thr Met His His Leu Leu Leu Glu Val
56           80           85           90
58 gaa gtg ata gag ggc acc ctg cag tgc ccg gaa tct gga cgt atg ttc      341
59 Glu Val Ile Glu Gly Thr Leu Gln Cys Pro Glu Ser Gly Arg Met Phe
60           95          100          105
62 ccc atc agc cgc ggg atc ccc aac atg ctg ctg agt gaa gag gaa act      389
63 Pro Ile Ser Arg Gly Ile Pro Asn Met Leu Leu Ser Glu Glu Glu Thr

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/883,152

DATE: 04/08/2003

TIME: 14:47:35

Input Set : A:\Pto.amc

Output Set: N:\CRF4\04082003\I883152.raw

```

64          110          115          120
66 gag agt t gattgtgccca ggcgccagtt tttcttggtta tgactgtgta tttttgttga      446
67 Glu Ser
68          125
70 tctataccct gtttccgaat tctgccgtgt gtatccccaa cccttgaccc aatgacacca      506
71 aacacagtggt ttttgagctc ggtattatat atttttttct cattaaaggt ttaaaacc      564
73 <210> SEQ ID NO: 2
74 <211> LENGTH: 125
75 <212> TYPE: PRT
76 <213> ORGANISM: H. sapiens
78 <400> SEQUENCE: 2
79 Met Lys Leu Leu Thr His Asn Leu Leu Ser Ser His Val Arg Gly Val
80 1 5 10 15
81 Gly Ser Arg Gly Phe Pro Leu Arg Leu Gln Ala Thr Glu Val Arg Ile
82 20 25 30
83 Cys Pro Val Glu Phe Asn Pro Asn Phe Val Ala Arg Met Ile Pro Lys
84 35 40 45
85 Val Glu Trp Ser Ala Phe Leu Glu Ala Ala Asp Asn Leu Arg Leu Ile
86 50 55 60
87 Gln Val Pro Lys Gly Pro Val Glu Gly Tyr Glu Asn Glu Glu Phe
88 65 70 75 80
89 Leu Arg Thr Met His His Leu Leu Leu Glu Val Glu Val Ile Glu Gly
90 85 90 95
91 Thr Leu Gln Cys Pro Glu Ser Gly Arg Met Phe Pro Ile Ser Arg Gly
92 100 105 110
93 Ile Pro Asn Met Leu Leu Ser Glu Glu Glu Thr Glu Ser
94 115 120 125
95 <210> SEQ ID NO: 3
96 <211> LENGTH: 919
97 <212> TYPE: DNA
98 <213> ORGANISM: H. sapiens
100 <220> FEATURE:
101 <221> NAME/KEY: CDS
102 <222> LOCATION: (219)...(693)
104 <400> SEQUENCE: 3
105 tggcacgagg tggcacgagg gtccgggtcg ttgaggatta ggctgctcgg gcgtaaccgg      60
106 agctggggcg cggctgccaa gggccggccc ggaagtccca gcggtcttta aattctcccg      120
107 tgctagggcc agcctgcgca ttcttacctg tcggggtgcg gcgagtgtct cacctctctg      180
108 cacttccaag gactcttgtc atctgcctta ggcgggaa atg ctg ttg ctg gat tgc      236
109 Met Leu Leu Leu Asp Cys
110 1 5
112 aac ccc gag gtg gat ggt ctg aag cat ttg ctg gag aca ggg gcc tcg      284
113 Asn Pro Glu Val Asp Gly Leu Lys His Leu Leu Glu Thr Gly Ala Ser
114 10 15 20
116 gtc aac gca ccc ccg gat ccc tgc aag cag tcg cct gtc cac tta gcc      332
117 Val Asn Ala Pro Pro Asp Pro Cys Lys Gln Ser Pro Val His Leu Ala
118 25 30 35
120 gca gga agc ggc ctt gct tgc ttt ctt ctc tgg cag ctg caa acg ggc      380
121 Ala Gly Ser Gly Leu Ala Cys Phe Leu Leu Trp Gln Leu Gln Thr Gly

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/883,152

DATE: 04/08/2003

TIME: 14:47:35

Input Set : A:\Pto.amc

Output Set: N:\CRF4\04082003\I883152.raw

```

122      40      45      50
124 gct gac ctc aac cag cag gat gtt tta gga gaa gct cca cta cac aag      428
125 Ala Asp Leu Asn Gln Gln Asp Val Leu Gly Glu Ala Pro Leu His Lys
126 55      60      65      70
128 gca gca aaa gtt gga agc ctg gag tgc cta agc ctg ctt gta gcc agt      476
129 Ala Ala Lys Val Gly Ser Leu Glu Cys Leu Ser Leu Leu Val Ala Ser
130      75      80      85
132 gat gcc caa att gat tta tgt aat aag aac ggg caa aca gct gaa gat      524
133 Asp Ala Gln Ile Asp Leu Cys Asn Lys Asn Gly Gln Thr Ala Glu Asp
134      90      95      100
136 ctc gct tgg tca tgt gga ttt cca gac tgt gcc aag ttt ctt aca aca      572
137 Leu Ala Trp Ser Cys Gly Phe Pro Asp Cys Ala Lys Phe Leu Thr Thr
138      105      110      115
140 att aaa tgt atg cag aca ata aaa gca agt gaa cac cct gac agg aat      620
141 Ile Lys Cys Met Gln Thr Ile Lys Ala Ser Glu His Pro Asp Arg Asn
142      120      125      130
144 gat tgt gtt gcc gtg ctc aga cag aaa cgg agt ctc gga agt gta gaa      668
145 Asp Cys Val Ala Val Leu Arg Gln Lys Arg Ser Leu Gly Ser Val Glu
146 135      140      145      150
148 aat acc agt ggg aaa agg aag tgc t gatgtcacgt gggttatgaa      713
149 Asn Thr Ser Gly Lys Arg Lys Cys
150      155
152 gaagtctgaa gaacgccttc atttcatgca aatctataag ctccctgcttt tggctttacc      773
153 atatgttggtg tctaattctcc ttctgagaag gacgaaaaac tttcttccaa gtgaagatcc      833
154 atttaagaac acatgtatatt acatgcctat aatatgctgg ttgtgtatgc tttgtctttt      893
155 aagttatttaa aggaacgtct aaaaaa      919
157 <210> SEQ ID NO: 4
158 <211> LENGTH: 158
159 <212> TYPE: PRT
160 <213> ORGANISM: H. sapiens
162 <400> SEQUENCE: 4
163 Met Leu Leu Leu Asp Cys Asn Pro Glu Val Asp Gly Leu Lys His Leu
164 1      5      10      15
165 Leu Glu Thr Gly Ala Ser Val Asn Ala Pro Pro Asp Pro Cys Lys Gln
166      20      25      30
167 Ser Pro Val His Leu Ala Ala Gly Ser Gly Leu Ala Cys Phe Leu Leu
168      35      40      45
169 Trp Gln Leu Gln Thr Gly Ala Asp Leu Asn Gln Gln Asp Val Leu Gly
170      50      55      60
171 Glu Ala Pro Leu His Lys Ala Ala Lys Val Gly Ser Leu Glu Cys Leu
172 65      70      75      80
173 Ser Leu Leu Val Ala Ser Asp Ala Gln Ile Asp Leu Cys Asn Lys Asn
174      85      90      95
175 Gly Gln Thr Ala Glu Asp Leu Ala Trp Ser Cys Gly Phe Pro Asp Cys
176      100      105      110
177 Ala Lys Phe Leu Thr Thr Ile Lys Cys Met Gln Thr Ile Lys Ala Ser
178      115      120      125
179 Glu His Pro Asp Arg Asn Asp Cys Val Ala Val Leu Arg Gln Lys Arg
180      130      135      140

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/883,152

DATE: 04/08/2003

TIME: 14:47:35

Input Set : A:\Pto.amc

Output Set: N:\CRF4\04082003\I883152.raw

```

181 Ser Leu Gly Ser Val Glu Asn Thr Ser Gly Lys Arg Lys Cys
182 145                      150                      155
183 <210> SEQ ID NO: 5
184 <211> LENGTH: 1949
185 <212> TYPE: DNA
186 <213> ORGANISM: H. sapiens
188 <220> FEATURE:
189 <221> NAME/KEY: CDS
190 <222> LOCATION: (5)...(1760)
192 <400> SEQUENCE: 5
193 caac atg gcg ccg tcc acg ccg ctc ttg aca gtc cga gga tca gaa gga      49
194 Met Ala Pro Ser Thr Pro Leu Leu Thr Val Arg Gly Ser Glu Gly
195 1 5 10 15
197 ctg tac atg gtg aat gga cca cca cat ttt aca gaa agc aca gtg ttt      97
198 Leu Tyr Met Val Asn Gly Pro Pro His Phe Thr Glu Ser Thr Val Phe
199 20 25 30
201 cca agg gaa tct ggg aag aat tgc aaa gtc tgt atc ttt agt aag gat      145
202 Pro Arg Glu Ser Gly Lys Asn Cys Lys Val Cys Ile Phe Ser Lys Asp
203 35 40 45
205 ggg acc ttg ttt gcc tgg ggc aat gga gaa aaa gta aat att atc agt      193
206 Gly Thr Leu Phe Ala Trp Gly Asn Gly Glu Lys Val Asn Ile Ile Ser
207 50 55 60
209 gtc act aac aag gga cta ctg cac tcc ttc gac ctc ctg aag gca gtt      241
210 Val Thr Asn Lys Gly Leu Leu His Ser Phe Asp Leu Leu Lys Ala Val
211 65 70 75
213 tgc ctt gaa ttc tca ccc aaa aat act gtc ctg gca acg tgg cag cct      289
214 Cys Leu Glu Phe Ser Pro Lys Asn Thr Val Leu Ala Thr Trp Gln Pro
215 80 85 90 95
217 tac act act tct aaa gat ggc aca gct ggg ata ccc aac cta caa ctt      337
218 Tyr Thr Thr Ser Lys Asp Gly Thr Ala Gly Ile Pro Asn Leu Gln Leu
219 100 105 110
221 tat gat gtg aaa act ggg aca tgt ttg aaa tct ttc atc cag aaa aaa      385
222 Tyr Asp Val Lys Thr Gly Thr Cys Leu Lys Ser Phe Ile Gln Lys Lys
223 115 120 125
225 atg caa aat tgg tgt cca tcc tgg tca gaa gat gaa act ctt tgt gcc      433
226 Met Gln Asn Trp Cys Pro Ser Trp Ser Glu Asp Glu Thr Leu Cys Ala
227 130 135 140
229 cgc aat gtt aac aat gaa gtt cac ttc ttt gaa aac aac aat ttt aac      481
230 Arg Asn Val Asn Asn Glu Val His Phe Phe Glu Asn Asn Asn Phe Asn
231 145 150 155
233 aca att gca aat aaa ttg cat ttg caa aaa att aat gac ttt gta tta      529
234 Thr Ile Ala Asn Lys Leu His Leu Gln Lys Ile Asn Asp Phe Val Leu
235 160 165 170 175
237 tca cct gga ccc caa cca tac aag gtg gct gtc tat gtt cca gga agt      577
238 Ser Pro Gly Pro Gln Pro Tyr Lys Val Ala Val Tyr Val Pro Gly Ser
239 180 185 190
241 aaa ggt gca cct tca ttt gtt aga tta tat cag tac ccc aac ttt gct      625
242 Lys Gly Ala Pro Ser Phe Val Arg Leu Tyr Gln Tyr Pro Asn Phe Ala
243 195 200 205

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/883,152

DATE: 04/08/2003

TIME: 14:47:35

Input Set : A:\Pto.amc

Output Set: N:\CRF4\04082003\I883152.raw

245	gga	cct	cat	gca	gct	tta	gct	aat	aaa	agt	ttc	ttt	aag	gca	gat	aaa	673
246	Gly	Pro	His	Ala	Ala	Leu	Ala	Asn	Lys	Ser	Phe	Phe	Lys	Ala	Asp	Lys	
247			210					215					220				
249	gtt	aca	atg	ctg	tgg	aat	aaa	gct	act	gct	gtg	ttg	gta	ata	gct		721
250	Val	Thr	Met	Leu	Trp	Asn	Lys	Lys	Ala	Thr	Ala	Val	Leu	Val	Ile	Ala	
251		225					230					235					
253	agc	aca	gat	gtt	gac	aag	aca	gga	gct	tcc	tac	tat	gga	gaa	caa	act	769
254	Ser	Thr	Asp	Val	Asp	Lys	Thr	Gly	Ala	Ser	Tyr	Tyr	Gly	Glu	Gln	Thr	
255	240					245					250				255		
257	cta	cac	tac	att	gca	aca	aat	gga	gaa	agt	gct	gta	gtg	caa	tta	cca	817
258	Leu	His	Tyr	Ile	Ala	Thr	Asn	Gly	Glu	Ser	Ala	Val	Val	Gln	Leu	Pro	
259				260						265				270			
261	aaa	aat	ggc	ccc	att	tat	gat	gta	gtt	tgg	aat	tct	agt	tct	act	gag	865
262	Lys	Asn	Gly	Pro	Ile	Tyr	Asp	Val	Val	Trp	Asn	Ser	Ser	Ser	Thr	Glu	
263			275					280					285				
265	ttt	tgt	gct	gta	tat	ggg	ttt	atg	cct	gcc	aaa	gcg	aca	att	ttc	aac	913
266	Phe	Cys	Ala	Val	Tyr	Gly	Phe	Met	Pro	Ala	Lys	Ala	Thr	Ile	Phe	Asn	
267			290					295					300				
269	ttg	aaa	tgt	gat	cct	gta	ttt	gac	ttt	gga	act	ggg	cct	cgt	aat	gca	961
270	Leu	Lys	Cys	Asp	Pro	Val	Phe	Asp	Phe	Gly	Thr	Gly	Pro	Arg	Asn	Ala	
271		305				310						315					
273	gcc	tac	tat	agc	cct	cat	gga	cat	ata	tta	gta	gct	gga	ttt	gga		1009
274	Ala	Tyr	Tyr	Ser	Pro	His	Gly	His	Ile	Leu	Val	Leu	Ala	Gly	Phe	Gly	
275	320				325					330				335			
277	aat	ctg	agg	gga	caa	atg	gaa	gtg	tgg	gat	gtg	aaa	aac	tac	aaa	ctt	1057
278	Asn	Leu	Arg	Gly	Gln	Met	Glu	Val	Trp	Asp	Val	Lys	Asn	Tyr	Lys	Leu	
279				340					345				350				
281	att	tct	aaa	ccg	gtg	gct	tct	gat	tct	aca	tat	ttt	gct	tgg	tgc	ccg	1105
282	Ile	Ser	Lys	Pro	Val	Ala	Ser	Asp	Ser	Thr	Tyr	Phe	Ala	Trp	Cys	Pro	
283			355					360					365				
285	gat	ggg	gag	cat	att	tta	aca	gct	aca	tgt	gct	ccc	agg	tta	cgg	gtt	1153
286	Asp	Gly	Glu	His	Ile	Leu	Thr	Ala	Thr	Cys	Ala	Pro	Arg	Leu	Arg	Val	
287		370				375						380					
289	aat	aat	gga	tac	aaa	att	tgg	cat	tat	act	ggc	tct	atc	ttg	cac	aag	1201
290	Asn	Asn	Gly	Tyr	Lys	Ile	Trp	His	Tyr	Thr	Gly	Ser	Ile	Leu	His	Lys	
291		385				390					395						
293	tat	gat	gtg	cca	tca	aat	gca	gaa	tta	tgg	cag	gtt	tct	tgg	cag	cca	1249
294	Tyr	Asp	Val	Pro	Ser	Asn	Ala	Glu	Leu	Trp	Gln	Val	Ser	Trp	Gln	Pro	
295	400				405						410			415			
297	ttt	ttg	gat	gga	ata	ttt	cca	gca	aaa	aca	ata	act	tac	caa	gca	gtt	1297
298	Phe	Leu	Asp	Gly	Ile	Phe	Pro	Ala	Lys	Thr	Ile	Thr	Tyr	Gln	Ala	Val	
299				420					425				430				
301	cca	agt	gaa	gta	ccc	aat	gag	gaa	cct	aaa	gtt	gca	aca	gct	tat	aga	1345
302	Pro	Ser	Glu	Val	Pro	Asn	Glu	Glu	Pro	Lys	Val	Ala	Thr	Ala	Tyr	Arg	
303				435					440				445				
305	ccc	cca	gct	tta	aga	aat	aaa	cca	atc	acc	aat	tcc	aaa	ttg	cat	gaa	1393
306	Pro	Pro	Ala	Leu	Arg	Asn	Lys	Pro	Ile	Thr	Asn	Ser	Lys	Leu	His	Glu	
307			450				455					460					
309	gag	gaa	cca	cct	cag	aat	atg	aaa	cca	caa	tca	gga	aac	gat	aag	cca	1441

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/883,152

DATE: 04/08/2003
TIME: 14:47:36

Input Set : A:\Pto.amc
Output Set: N:\CRF4\04082003\I883152.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:12; N Pos. 1047,1301,1507

Seq#:15; N Pos. 1047,1301,1507